

TECHNICAL NOTE

Open Access



# Optimizing a PCR protocol for *cpn60*-based microbiome profiling of samples variously contaminated with host genomic DNA

Lisa A Johnson<sup>1</sup>, Bonnie Chaban<sup>1,2</sup>, John C S Harding<sup>3</sup> and Janet E Hill<sup>1\*</sup>

## Abstract

**Background:** The current recommended protocol for chaperonin-60 (*cpn60*) universal target based microbiome profiling includes universal PCR of microbiome samples across an annealing temperature gradient to maximize the diversity of sequences amplified. However, the value of including this gradient approach has not been formally evaluated since the optimization of a modified universal PCR primer cocktail for *cpn60* PCR. PCR conditions that maximize representation of the microbiome while minimizing PCR-associated distortion of the community structure, especially in samples containing large amounts of host genomic DNA are critical. The goal of this study was to measure the effects of PCR annealing temperature and the ratio of host to bacterial DNA on the outcome of microbiota analysis, using pig microbiota as a model environment.

**Findings:** Six samples were chosen with an anticipated range of ratios of pig to bacterial genomic DNA, and universal *cpn60* PCR amplification with an annealing temperature gradient was used to create libraries for pyrosequencing, resulting in 426,477 sequences from the six samples. The sequences obtained were classified as target (*cpn60*) or non-target based on the percent identity of their closest match to the cpnDB reference database, and target sequences were further processed to create microbiome profiles for each sample at each annealing temperature. Annealing temperature affected the amount of PCR product generated, with more product generated at higher temperatures. Samples containing proportionally more host genomic DNA yielded more non-target reads, especially at lower annealing temperatures. However, microbiome composition for each sample across the annealing temperature gradient remained consistent at both the phylum and operational taxonomic unit levels. Although some microbial sequences were detected at only one annealing temperature, these sequences accounted for a minority of the total microbiome.

**Conclusions:** These results indicate that PCR annealing temperature does have an affect on *cpn60* based microbiome profiles, but that most of the differences are due to differences in detection of low abundance sequences. Higher annealing temperatures resulted in larger amounts of PCR product and lower amounts of non-target sequence amplification, especially in samples containing proportionally large amounts of host DNA. Taken together these results provide important information to guide decisions about experimental design for *cpn60* based microbiome studies.

**Keywords:** *cpn60*, Microbiome profiling, Pig microbiome, PCR conditions, Gradient PCR, Pyrosequencing, Host microbiome

\*Correspondence: janet.hill@usask.ca

<sup>1</sup> Department of Veterinary Microbiology, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, SK S7N 5B4, Canada  
Full list of author information is available at the end of the article

## Background

Complex microbial communities play a fundamental role in the health of animals, humans and the environment, and new opportunities to more fully characterize and understand these communities are available due to advances in sequencing technologies. The chaperonin-60 universal target (*cpn60* UT, a 549–567 bp segment of the *cpn60* gene) has been demonstrated to be a preferred barcode sequence for bacteria [1], and an especially useful target for high resolution microbiome profiling [2–10]. However, this approach, like any other that relies on PCR amplification of a target gene, distorts our view of the true community composition [11, 12].

One parameter of PCR that contributes to amplification bias is the annealing temperature. Previous work has indicated that performing the *cpn60* UT PCR across a range of annealing temperatures and pooling the results improves the comprehensive assessment of microbial diversity of a sample [13]. The current recommended protocol involves 12 PCR reactions across an annealing temperature gradient for each sample that are subsequently pooled to create a sequencing library [14]. Including a full temperature gradient adds expense and logistical complexity to the protocol, and the importance of this procedure has not been re-evaluated since the development and application of a universal primer cocktail that improves representation of high G + C content organisms [15, 16], and coincident adoption of new generations of Taq polymerases and other PCR reagents.

Another parameter that may affect how a microbial community is amplified is the ratio of host to microbial genomic DNA. Large proportions of host genomic DNA can interfere with the amplification of *cpn60* UT, resulting in non-target amplification. A previous study of the upper respiratory tract of pandemic H1N1 influenza patients demonstrated that samples containing large proportions of human DNA resulted in non-target amplification originating from the human genome during sequencing [3]. Although comprehensive microbiome profiles were generated, it was with additional, perhaps unnecessary cost due to greater sequencing depth requirements to ensure adequate coverage of the microbiome in the context of large amounts of host genomic DNA.

The goal of this study was to measure the effects of PCR annealing temperature and the ratio of host to bacterial DNA on the outcome of *cpn60*-based microbiota analysis using the recommended PCR primer cocktail, and pig microbiota as a model environment.

## Methods

### Sample collection and DNA extraction

Six samples were collected, five from different anatomical sites within a single pig, and one from an environmental

soil sample. Samples were chosen to represent an anticipated range of pig to bacterial DNA ratios, from entirely pig to entirely bacterial. Samples were collected from a pig with congenital nasal obstruction, euthanized in an unrelated study designed and conducted in accordance with the Canadian Council for Animal Care and approved by the University of Saskatchewan Committee on Animal Care and Supply (Protocol #20130034). Samples collected included a section of brain tissue, a nasal swab, a stomach mucosal scraping, a colon mucosal scraping and feces (collected from rectum). Soil collected from the grounds of the University of Saskatchewan campus, was used as a pig genomic DNA free sample. All samples were stored at  $-20^{\circ}\text{C}$  until DNA extraction.

Total genomic DNA extraction was performed on 0.2 g of material from the six samples using a previously described protocol that combines mechanical disruption and chemical extraction [17]. An extraction blank, consisting of ultrapure water, was carried through DNA extraction and sequencing. The DNA concentration of each sample extract was measured by spectrophotometry, and samples were diluted to a concentration of 10 ng/ $\mu\text{l}$  for PCR.

### Quantitative PCR (qPCR)

Quantification of bacterial and pig genomic DNA in each sample was performed by qPCR using the primers as described in Table 1. Estimation of bacterial content was obtained using primers targeting the V1–V3 region the 16S rRNA gene [18] in a qPCR assay described previously [3], while estimation of pig genomic DNA content was based on the quantification using the pig alpha actin gene.

All qPCR reactions were performed in duplicate, including a no template control (NTC) and a standard curve, which was prepared with the target-containing plasmids at concentrations of  $10^0$ – $10^7$  copies/reaction. Each reaction contained  $1 \times$  iQ SYBR Green Supermix (Bio-Rad, Mississauga, ON, Canada), 400 nM of both forward and reverse primers, and 2  $\mu\text{l}$  of template DNA, in a final volume of 25  $\mu\text{l}$ . All reactions used the following program:  $95^{\circ}\text{C}$  for 3 min, 40 cycles of  $95^{\circ}\text{C}$  for 15 s, annealing temperature (Table 1) for 15 s,  $72^{\circ}\text{C}$  for 15 s, and final extension at  $72^{\circ}\text{C}$  for 5 min. A dissociation curve was also run for 81 cycles at  $0.5^{\circ}\text{C}$  increments from  $55$  to  $95^{\circ}\text{C}$ . PCR was performed on a MyiQ thermocycler (Bio-Rad), and the data was analyzed using iQ5 Optical System Software.

### *cpn60* universal target PCR and pyrosequencing

Amplification of the UT region of the *cpn60* gene was performed using a primer cocktail consisting of a 1:3 molar ratio of primers H279/H280:H1612/H1613 [15]

**Table 1 PCR primer sequences**

Target	Primer	Sequence (5'–3') <sup>a</sup>	Annealing temp (°C)	Reference
cpn60 UT	H279	GAIIIGCIGGIGAYGGIACIAC	42, 48, 54, 60	[15]
	H280	YKIYKITCICCRAAICIGGICYTT		
	H1612	GAIIIGCIGGYGACGGYACSACSAC		
	H1613	CGRCGRTCRCGGAAGCCSGGIGCCTT		
16S rRNA gene	SRV3-1	CGGYCCAGACTCTAC	62	[18]
	SRV3-2	TTACCGCGGCTGCTGGCAC		
Pig $\alpha$ -actin gene	JH0462	CCCAGAGCAAGCGAGGTATT	68	This study
	JH0463	GGGCCTCAGTGAGCAGAGTA		

<sup>a</sup> I = inosine, Y = G or T, K = G or T, R = A or G, S = G or C.

(Table 1). Primers were modified with the addition of multiplexing ID tags at the 5' end. Each PCR reaction contained 1 × PCR reaction buffer, 2.5 mM MgCl<sub>2</sub>, 200  $\mu$ M dNTP, 800 nM primer cocktail, 2.5 U Platinum Taq DNA Polymerase (Invitrogen, Burlington, ON, Canada) and 2  $\mu$ l of template DNA, in a final volume of 50  $\mu$ l. For each sample, four PCR reactions were performed in a thermocycler (Eppendorf Mastercycler) over an annealing temperature gradient with the following program: 94°C for 5 min, 40 cycles of 95°C for 30 s, annealing temperature (Table 1) for 30 s, and a final extension of 72°C for 2 min. A NTC and positive amplification control were included with each gradient. PCR products (5  $\mu$ l) were resolved and visualized on a 1% agarose gel by ethidium bromide staining. Products were then purified by agarose gel extraction (QIAEX II gel extraction kit, Qiagen, Mississauga, ON, Canada) and suspended in TE buffer [10 mM Tris (pH 8), 1 mM EDTA]. The resulting 28 amplicon libraries were prepared and pooled in equimolar concentrations and sequenced using the Roche 454 GS Junior system as per manufacturer's instructions.

#### OTU assembly and analysis

Raw pyrosequencing data was processed by using the default on-rig procedures from 454/Roche. Filter-passing reads were used in the subsequent analyses for each of the pyrosequencing libraries. Sequence reads were demultiplexed and processed to generate operational taxonomic units (OTU) with the microbial Profiling Using Metagenomic Assembly (mPUMA) pipeline [19] using Trinity for OTU assembly. Processing of sequence reads by mPUMA includes identification and removal of amplification primer sequences and identification of putative chimeras using the C3 chimera checker. Watered-BLAST [20] comparison to the cpnDB\_nr reference database (version 20130321, downloaded from <http://www.cpnadb.ca>) [21] was used to identify each OTU. Sequences identified as non-target were reference mapped (GS Reference Mapper, Roche, Bradford, CT, USA) to the pig genome

(*Sus scrofa*, Genbank Accession AEMK01000000) to determine the amount of non-target amplification of pig genome origin. Coverage and Shannon diversity for each library was calculated using Mothur [22]. Principal coordinates analysis of jackknifed Bray–Curtis dissimilarity matrices was performed in QIIME [23].

## Results and discussion

### Genomic DNA extraction and quantification

Estimated copy numbers of pig genomes and bacterial genomes were determined by qPCR targeting the pig  $\alpha$ -actin or bacterial 16S rRNA genes. Starting quantities were calculated based on interpolation using a standard curve of a ten-fold dilution series of plasmid containing the target sequence. For the  $\alpha$ -actin PCR, a linear standard curve was obtained over a range of 10<sup>2</sup>–10<sup>7</sup> copies per reaction (efficiency 105%,  $r^2 = 0.996$ ). For the 16S rRNA PCR, a linear standard curve was obtained over a range of 10<sup>4</sup>–10<sup>7</sup> copies per reaction (efficiency 73%,  $r^2 = 0.955$ ). Based on the estimate copy numbers, ratios of pig and bacterial genomic DNA were 80:1 (brain), 0.8:1 (nasal), 2.3:1 (stomach), 1.3:1 (colon), 0.0004:1 (feces), and 0:100 (soil) (Table 2).

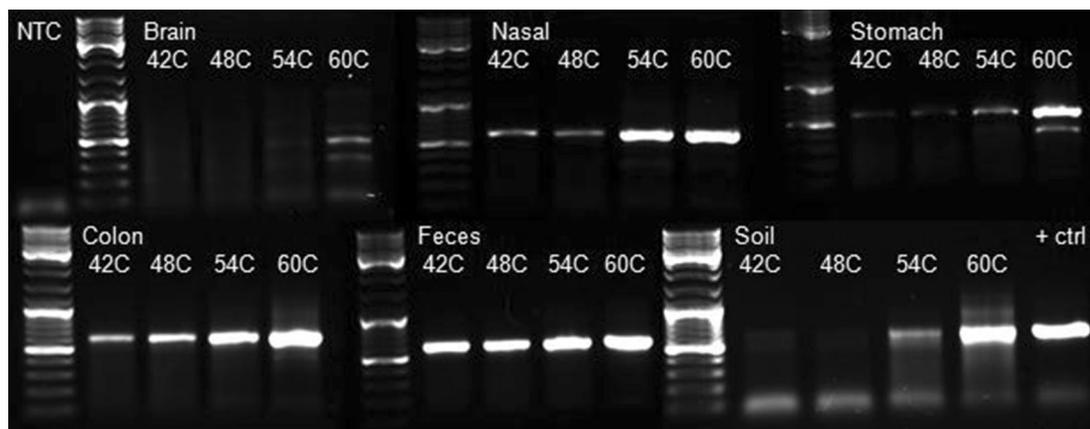
### cpn60 PCR and pyrosequencing

The cpn60 UT PCR products generated for each sample across the annealing temperature gradient were visualized on an agarose gel (Figure 1). The amount of PCR product produced varied by sample type and annealing temperature. For all sample types, the amount of PCR product generated increased with higher annealing temperatures, indicating that annealing temperature can have a dramatic effect on the amount of PCR product produced. The relationship between annealing temperature and PCR efficiency is well known [24]. Increasing hybridization stringency by elevating the annealing temperature reduces the occurrence of non-productive annealing events where primers anneal to non-target regions of the template. Although these interactions could lead to primer extension, the resulting products

**Table 2 Quantification of genomic DNA extracted from pig tissue and soil microbiomes**

Sample	Gene copies per 20 ng DNA <sup>a</sup>		
	Pig $\alpha$ -actin	Bacterial 16S rRNA	Pig:bacteria DNA ratio
Brain	$9.57 \times 10^5 \pm 2.75 \times 10^5$	$1.20 \times 10^4 \pm 1.79 \times 10^3$	80:1
Nasal cavity	$5.65 \times 10^6 \pm 7.07 \times 10^4$	$6.77 \times 10^6 \pm 6.91 \times 10^5$	0.8:1
Stomach mucosa	$4.76 \times 10^6 \pm 1.23 \times 10^6$	$2.04 \times 10^6 \pm 1.26 \times 10^5$	2.3:1
Colon mucosa	$8.21 \times 10^6 \pm 2.58 \times 10^3$	$6.08 \times 10^6 \pm 3.30 \times 10^4$	1.3:1
Feces	$1.09 \times 10^5 \pm 6.21 \times 10^3$	$2.45 \times 10^8 \pm 4.79 \times 10^7$	0.0004:1
Soil	Not detected	$3.81 \times 10^5 \pm 2.82 \times 10^5$	All bacterial

<sup>a</sup> Quantities reported are the average of duplicates  $\pm$  standard deviation.



**Figure 1** *cpn60* PCR products from each sampling site at each PCR annealing temperature tested visualized on a 1% agarose gel. An equal volume of PCR product (5  $\mu$ l) from each sampling site (brain, nasal, stomach, colon, feces, and soil) and each PCR annealing temperature gradient point tested (42, 48, 54 and 60°C) were visualized next to a ladder to determine amplification efficiency. Major bands on the DNA ladder indicate 500, 1,000 and 3,000 bp, while the *cpn60* UT product is ~650 bp. NTC no template control, +ctrl positive control sample (genomic DNA from *Helicobacter canis*).

would be unlikely to participate in further amplification and could interfere with amplification of target regions by providing primer annealing sites.

Sequencing from the 28 libraries resulted in 426,477 high quality reads, with library sizes ranging from 1,707 to 81,128 reads (median 10,124). To facilitate comparison and reduce errors in interpretation due to unequal sampling depth [25], libraries were sub-sampled to the size of the smallest library of 1,707 reads. Coverage values for the subsampled pig derived amplicon libraries were >0.94. Coverage for the soil libraries ranged from 0.81 to 0.88, reflecting the greater diversity of these samples (see below). Reads were assembled into OTU and compared to the cpnDB reference database for assignment of nearest neighbour taxonomic labels to *cpn60* (target) OTU, and to identify non-target sequences. OTU sequences were classified as target (*cpn60*) or non-target based on the percent identity of their closest database match. Our previous experience and analysis of *cpn60*

reference data in cpnDB has led the establishment of a cutoff of 55% identity for discrimination of *cpn60* and non-target sequences. Additional file 1 shows the typical, bimodal distribution of percent identity values observed when amplicon libraries include non-target sequences. Detailed analysis of non-target OTU is described below. OTU identified as *cpn60* sequences (1164/2164 OTU assembled) corresponded to 480 nearest neighbour species: 380 OTU detected in soil, 104 in colon, 103 in feces, 52 in stomach, 23 in nasal swab, and 36 in brain. A summary of the nearest neighbour species detected in each sample and the number of sequence reads associated with each is given in Additional file 2.

#### Target and non-target amplification

OTU with <55% identity to the closest match in the cpnDB reference database were considered non-target and removed from microbiota analysis. Sequence reads contributing to these non-target OTU were screened

against the pig genome to identify non-specific amplification products of pig origin. Any remaining sequences, likely of microbial origin, were assigned to the category non-target (other origin). The proportions of *cpn60* target and non-target (pig and other origin) sequence reads in each library are shown in Figure 2. The average target:non-target ratio across the four annealing temperature libraries for each sample were 4:96 (brain), 98:2 (nasal), 94:6 (stomach), 96:4 (colon), 99:1 (feces) and 96:4 (soil). Furthermore, samples containing appreciable proportions of pig genomic DNA (brain, nasal swab, stomach and colon mucosal scrapings) yielded more non-target reads corresponding to the pig genome, especially at lower annealing temperatures. For example, the brain sample, which was estimated to contain 99% pig genomic DNA, resulted in 89% of the total reads corresponding to non-target pig genome amplification. Alternatively, the nasal and stomach samples, which contained 45 and 70% pig genomic DNA respectively, resulted in 1 and 4% of the total reads to correspond to the pig genome.

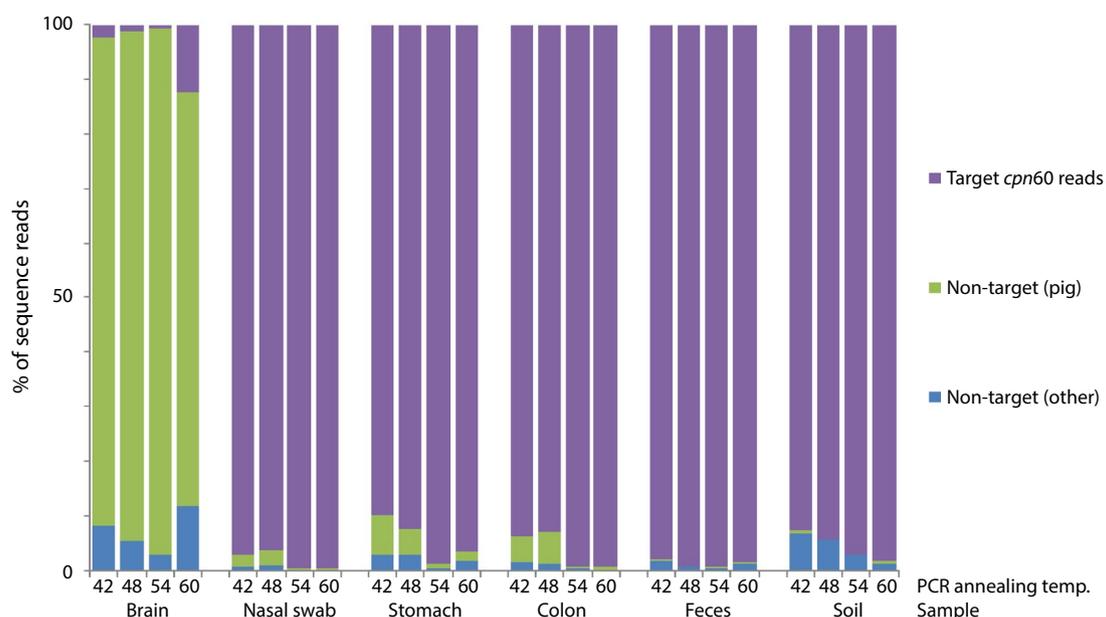
The profiles corresponding to the brain sample libraries contained from 0.8 to 12.3% bacterial *cpn60* sequences (269 reads total) (Figure 2). While some of these sequence reads may originate from the brain tissue, trace contamination was detected in the extraction negative and some of these sequences were also detected in the brain libraries. Thus, the detection of microbial sequences the brain libraries was at least partially due to contamination of the sample during collection at necropsy, and/or during

laboratory processing. Brain microbial reads were not analyzed further.

There was an observed trend of greater proportions of non-target amplification at lower annealing temperatures, especially in samples containing large proportions of host DNA (Figure 2). This finding is consistent with a previous *cpn60* microbiome study where non-target amplification from samples of the human upper respiratory tract, containing an order of magnitude more human DNA than bacterial DNA, accounted for up to 85% of the data generated [5].

#### Effect of annealing temperature on phylum level profiles

*cpn60* OTU were used to generate phylum level microbiome profiles, which were generally consistent within body site, regardless of annealing temperature (Figure 3). The nasal passage libraries were dominated by Proteobacteria (average of 4 annealing temperature libraries, 79%), Actinobacteria (13%) and Firmicutes (8%), similar to previous descriptions of the pig tonsil microbiome [26]. The stomach libraries contained mostly Firmicutes (91%), with some Actinobacteria (6%) and Proteobacteria (3%), consistent with the most abundant phyla identified in the stomach microbiomes of horses [27]. The colon and feces libraries were dominated by Firmicutes (45 and 53%, respectively), Proteobacteria (46 and 29%), and Bacteroidetes (9 and 18%) as expected based on previous descriptions of these environments [28, 29]. The soil libraries included Proteobacteria (37%), Actinobacteria



**Figure 2** Proportion of each pyrosequencing library that was composed of target (*cpn60*) reads and non-target (of pig or other origin) reads. Libraries are labeled by sample site and PCR annealing temperature.

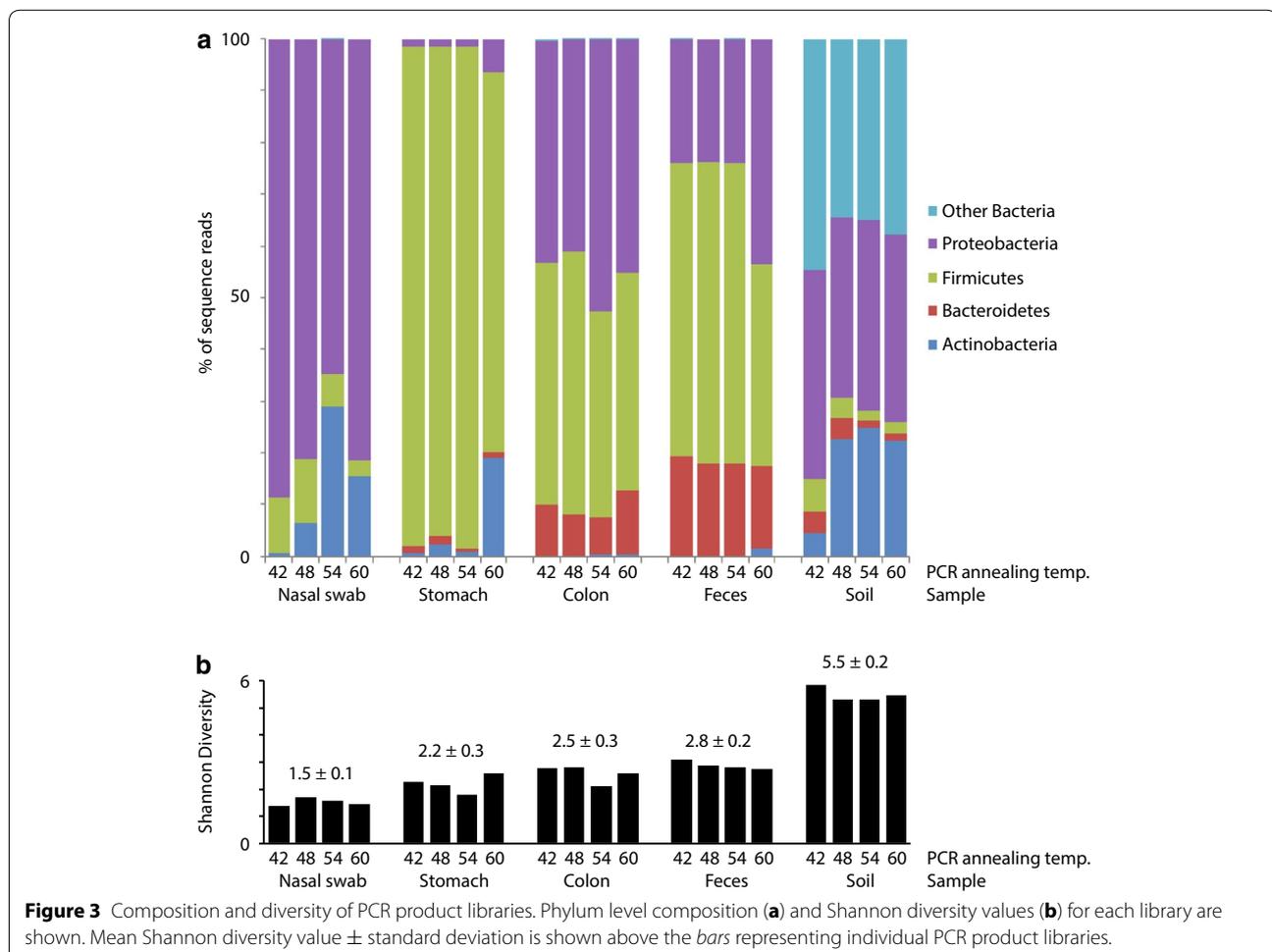
(19%), Bacteroidetes (3%), and Firmicutes (3%), with 38% of the reads belonging to other phyla. Diversity was lowest in the nasal swab libraries (Shannon diversity  $1.5 \pm 0.1$ ) and highest in the soil libraries ( $5.5 \pm 0.2$ ), and no consistent relationship of annealing temperature and library diversity was observed (Figure 3b).

#### Effect of annealing temperature on species level profiles

To expand on the phylum level comparisons and make a more detailed comparison of the microbial profiles produced across the annealing temperature gradient, we investigated whether different taxa were detected at different annealing temperatures. OTU sequences with the same nearest neighbour in cpnDB were combined as nearest neighbour “species” and their abundances combined accordingly for this analysis (Additional file 2). At this level, the overall similarity between microbiome profiles generated at different annealing temperatures is also apparent. Figure 4 shows the results of principal coordinates analysis based on Bray–Curtis dissimilarity values for species level microbiome profiles of all samples

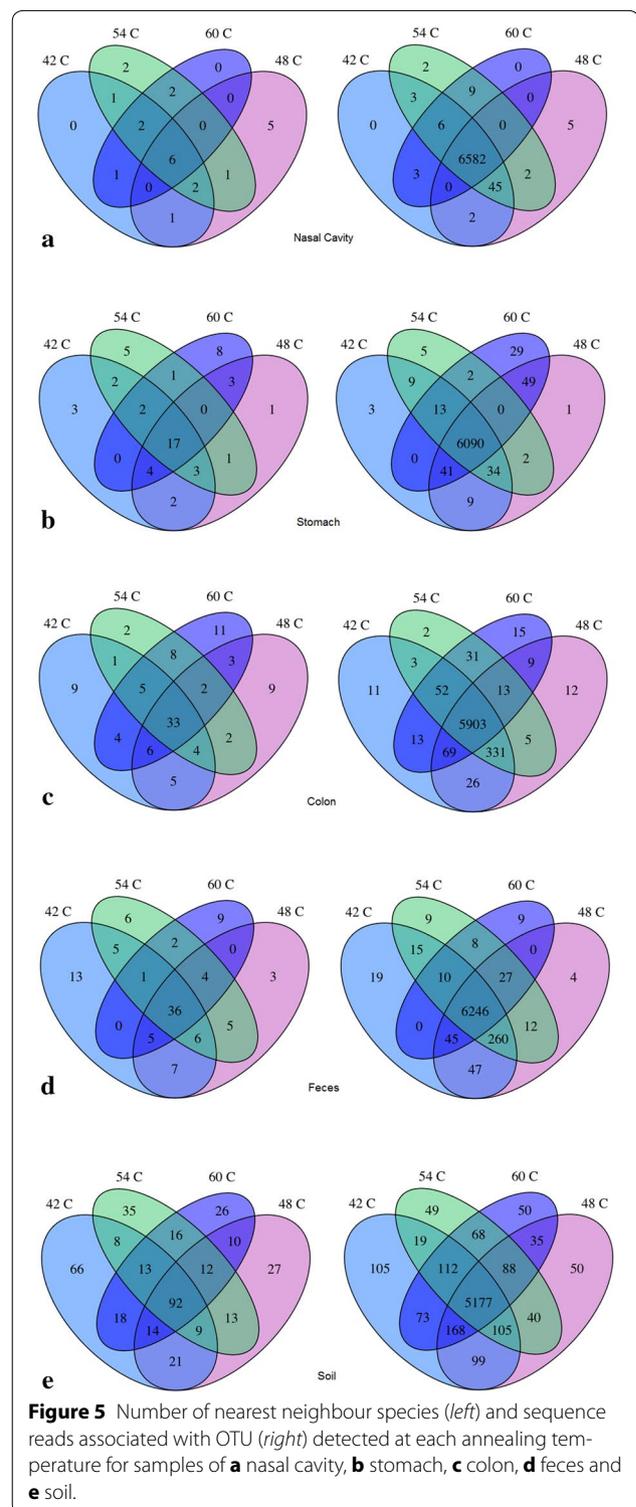
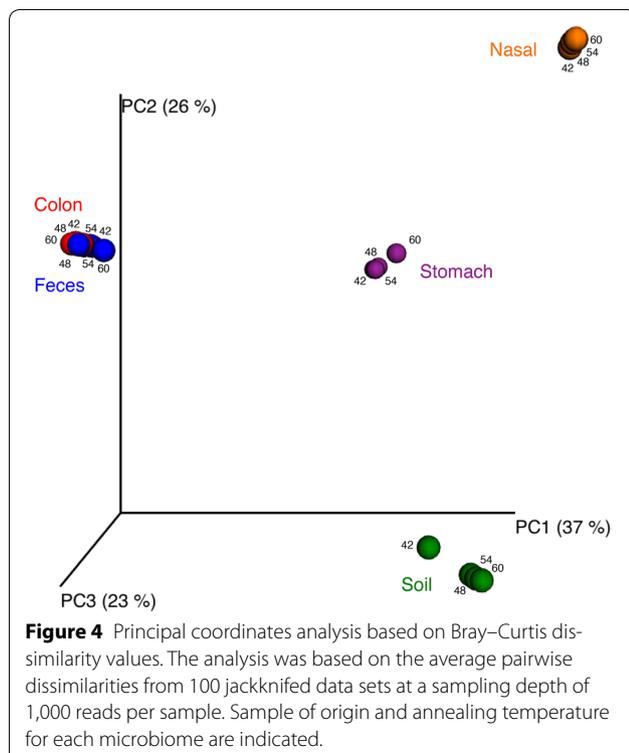
and annealing temperatures. Microbiome profiles cluster according to tissue or material of origin, with feces and colon overlapping.

To further understand the relationship between OTU abundance and prevalence in the different annealing temperature libraries we examined the distribution of individual OTU sequences. Figure 5 shows the number of nearest neighbour species (left side) and the number of sequence reads associated with OTU (right side) detected at specific annealing temperatures for each sample type. For each sample type, the total number of sequence reads was obtained by combining the sub-sampled libraries for each annealing temperature. The proportion of total species that were detected at all four annealing temperatures ranged from 24.2% (92/380) for soil to 35.3% (36/102) for feces (Figure 5, left). However, when abundance was considered, it was apparent that the majority of sequence reads were associated with OTU that were detected at all annealing temperatures (Figure 5, right). For all samples, OTU sequences detected at all annealing temperatures accounted for a



minimum of 83.0% of reads. For example, although 8 species were detected only in the 60°C stomach library, these OTU accounted for only 1.7% of the data. Overall, for the pig microbiomes analyzed, from 86.5% (colon) to 96.4% (nasal swab) of the sequence reads were associated with OTU detected at all annealing temperatures. Interestingly, the proportion of sequence reads associated with OTU detected at all annealing temperatures was noticeably lower for the soil library where these common OTU accounted for only 75.8% of the sequence data, which reflects the greater diversity and more even frequency distribution of OTU in this microbiome relative to the pig associated communities.

Nearest neighbour species could be detected at specific annealing temperatures because they are preferentially amplified at those annealing temperatures, however this was not addressed in our study. An alternative explanation is that these OTU were simply so rare in the starting sample that they were only detected sporadically at the sequencing depth used. For example, in the nasal libraries, 2/23 OTU were detected at 42, 54 and 60°C but not at 48°C (Figure 5a, left). However, these OTU account for only 6/6,828 sequence reads and so it seems plausible that additional sequencing depth in the 48°C library would result in their detection. Indeed, it was observed that for samples where additional sequence reads were available, addition of more sequences into the analysis



resulted in detection of more of the rare OTU in more annealing temperature libraries, and thus fewer nearest neighbours being detected only at one annealing temperature (data not shown).

## Conclusion

The universal *cpn60* PCR annealing temperature had effects on the amount of PCR product produced and the microbiota profiles produced at both phylum and OTU levels. However, the differences observed in profiles were largely due to low abundance sequences, especially in the animal associated microbiomes examined. The amount of non-target amplification was somewhat greater at lower annealing temperatures, but was most prominent for samples containing large proportions of host genomic DNA. OTU uniquely detected at certain annealing temperatures were rare, and accounted for a small proportion of microbiota profiles, although this finding was affected by sampling depth. Taken together, these results indicate that the logistically complex and expensive process of including a full temperature gradient in *cpn60* amplicon library production may not be necessary, depending on the goals of the experiment. In cases where samples are known to contain large amounts of host DNA, amplification at the maximum annealing temperature feasible will minimize non-target amplification. However, practitioners of *cpn60* based microbiome profiling should also consider that if detection of rare OTU is critical to the experimental goals, an annealing temperature gradient may still prove useful to increase the likelihood of capturing of all of the microbial diversity present in the environment of interest.

**Availability of supporting data** The data set supporting the results of this article is available in the NCBI SRA repository, and is associated with BioProject accession PRJNA260274, <http://www.ncbi.nlm.nih.gov/bioproject/PRJNA200951>.

## Additional files

**Additional File 1:** Title of data: Distribution of percent identities of *cpn60* OTU sequences to their nearest neighbour in the cpnDB reference database. Description of data: The lower distribution (OTU with <55% identity) contains non-*cpn60* sequences, largely pig genomic DNA sequences, which were subsequently mapped on to the pig genome for confirmation of their origin. The majority of pig genomic DNA sequences were assembled from the brain tissue sample, which also had the highest ratio of host:bacteria genomic DNA of any of the samples in the study. Manual inspection of OTU sequences with <55% identity but that did not map on to the pig genome indicated that many of them were of bacterial origin (bacterial, non-*cpn60* sequences).

**Additional File 2:** Title of data: Nearest neighbour species detection frequencies. Description of data: The number of sequence reads associated with each nearest neighbour species detected in each sample library are shown.

## Abbreviations

OTU: operational taxonomic unit; *cpn60*: chaperonin-60; UT: universal target; NTC: no template control; mPUMA: microbial Profiling Using Metagenomic Assembly

## Authors' contributions

JEH, JCSh and BC designed the experiment. LAJ and JCSh collected samples. LAJ performed the microbiome sample preparation and data generation. BC performed the qPCR. LAJ and JEH performed the microbiome analysis. LAJ drafted the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> Department of Veterinary Microbiology, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, SK S7N 5B4, Canada.

<sup>2</sup> Present Address: Department of Life Sciences, Imperial College London, London SW7 2AZ, UK. <sup>3</sup> Department of Large Animal Clinical Sciences, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, SK S7N 5B4, Canada.

## Acknowledgements

LAJ was supported by an Undergraduate Student Research Award from the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors wish to thank Champika Fernando for technical assistance.

## Compliance with ethical guidelines

## Competing interests

The authors declare that they have no competing interests.

Received: 3 September 2014 Accepted: 13 May 2015

Published online: 20 June 2015

## References

1. Links MG, Dumonceaux TJ, Hemmingsen SM, Hill JE (2012) The chaperonin-60 universal target is a barcode for bacteria that enables de novo assembly of metagenomic sequence data. *PLoS One* 7(11):e49755
2. Chaban B, Links MG, Paramel Jayaprakash T, Wagner EC, Bourque DK, Lohn Z et al (2014) Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. *Microbiome* 2:23
3. Chaban B, Albert A, Links MG, Gardy J, Tang P, Hill JE (2013) Characterization of the upper respiratory tract microbiomes of patients with pandemic H1N1 influenza. *PLoS One* 8(7):e69559
4. Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ (2014) Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytol* 202(2):542–553
5. Desai AR, Links MG, Collins SA, Mansfield GS, Drew MD, Van Kessel AG et al (2012) Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 350:134–142
6. Schellenberg JJ, Links MG, Hill JE, Dumonceaux TJ, Kimani J, Jaoko W et al (2011) Molecular definition of vaginal microbiota in East African commercial sex workers. *Appl Environ Microbiol* 77(12):4066–4074
7. Oliver KL, Hamelin RC, Hintz WE (2008) Effects of transgenic hybrid aspen overexpressing polyphenol oxidase on rhizosphere diversity. *Appl Environ Microbiol* 74(17):5340–5348
8. Pratt DL, Dumonceaux TJ, Links MG, Fonstad TA (2012) Influence of mass burial of animal carcasses on the types and quantities of microorganisms within a burial site. *Trans Am Soc Agric Biol Eng* 55(6):2195–2212
9. Town JR, Links MG, Fonstad TA, Dumonceaux TJ (2014) Molecular characterization of anaerobic digester microbial communities identifies microorganisms that correlate to reactor performance. *Bioresour Technol* 151:249–257
10. Town J, Annand H, Pratt D, Dumonceaux T, Fonstad T (2014) Microbial community composition is consistent across anaerobic digesters processing wheat-based fuel ethanol waste streams. *Bioresour Technol* 157:127–133
11. Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol* 64(10):3724–3730
12. Ishii K, Fukui M (2001) Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR. *Appl Environ Microbiol* 67(8):3753–3755

13. Hill JE, Seipp RP, Betts M, Hawkins L, Van Kessel AG, Crosby WL et al (2002) Extensive profiling of a complex microbial community by high-throughput sequencing. *Appl Environ Microbiol* 68(6):3055–3066
14. Schellenberg J, Links MG, Hill JE, Hemmingsen SM, Peters GA, Dumonceaux TJ (2011) Pyrosequencing of chaperonin-60 (*cpn60*) amplicons as a means of determining microbial community composition. *Methods Mol Biol* 733:143–158
15. Hill JE, Town JR, Hemmingsen SM (2006) Improved template representation in *cpn60* PCR product libraries generated from complex templates by application of a specific mixture of PCR primers. *Environ Microbiol* 8(4):741–746
16. Hill JE, Fernando WMU, Zello GA, Tyler RT, Dahl WJ, Van Kessel AG (2010) Improvement of the representation of bifidobacteria in fecal microbiota metagenomic libraries by application of the *cpn60* universal primer cocktail. *Appl Environ Microbiol* 76(13):4550–4552
17. Dumonceaux TJ, Hill JE, Briggs SA, Amoako KK, Hemmingsen SM, Van Kessel AG (2006) Enumeration of specific bacterial populations in complex intestinal communities using quantitative PCR based on the chaperonin-60 target. *J Microbiol Meth* 64(1):46–62
18. Lee DH, Zo YG, Kim SJ (1996) Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism. *Appl Environ Microbiol* 62(9):3112–3120
19. Links MG, Chaban B, Hemmingsen SM, Muirhead K, Hill JE (2013) mPUMA: a computational approach to microbiota analysis by de novo assembly of OTUs based on protein-coding barcode sequences. *Microbiome* 1:23
20. Schellenberg J, Links MG, Hill JE, Dumonceaux TJ, Peters GA, Tyler S et al (2009) Pyrosequencing of the chaperonin-60 universal target as a tool for determining microbial community composition. *Appl Environ Microbiol* 75(9):2889–2898
21. Hill JE, Penny SL, Crowell KG, Goh SH, Hemmingsen SM (2004) cpnDB: a chaperonin sequence database. *Genome Res* 14(8):1669–1675
22. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75(23):7537–7541
23. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336
24. Rychlik W, Spencer WJ, Rhoads RE (1990) Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Res* 18(21):6409–6412
25. Gihring TM, Green SJ, Schadt CW (2012) Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environ Microbiol* 14(2):285–290
26. Lowe BA, Marsh TL, Isaacs-Cosgrove N, Kirkwood RN, Kiupel M, Mulks MH (2012) Defining the “core microbiome” of the microbial communities in the tonsils of healthy pigs. *BMC Microbiol* 12:20
27. Perkins GA, den Bakker HC, Burton AJ, Erb HN, McDonough SP, McDonough PL et al (2012) Equine stomachs harbor an abundant and diverse mucosal microbiota. *Appl Environ Microbiol* 78(8):2522–2532
28. Hill JE, Seipp RP, Betts M, Hawkins L, Van Kessel AG (2002) Extensive profiling of a complex microbial community using genomic techniques. In: Canadian Society of Microbiologists 52nd Annual General Meeting. Saskatoon, Saskatchewan
29. Leser TD, Lindecrona RH, Jensen TK, Jensen BB, Moller K (2000) Changes in bacterial community structure in the colon of pigs fed different experimental diets and after infection with *Brachyspira hyodysenteriae*. *Appl Environ Microbiol* 66(8):3290–3296

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

